ESSAY 9.1
Moral Conclusions from Non-moral Premises

Stephen Maitzen

1. Introduction

Recent decades have seen a long-standing debate over the 'the logical autonomy of ethics' or whether it's possible to derive an 'ought' from an 'is'. Although the issue is traceable to Hume, in the last half century a Hume-independent controversy has raged regarding four key questions: Can we derive, by purely logical means, a substantively moral conclusion from substantively non-moral premises? If so, can such a derivation be deductively valid without being trivially valid? If valid, can it be sound? If sound, can it be ethically nontrivial? I answer 'yes' to all four questions. If I'm right, autonomists are wrong on all four counts.

2. The Dialectical Context

Anti-autonomists have tried a variety of ingenious strategies for deriving moral conclusions from non-moral premises, but none have convinced their opponents. My strategy differs from others I've seen employed. Although my use of the disjunctive syllogism resembles that of A. N. Prior, my strategy differs from Prior's in several important ways. I defend my derivations of 'ought' from 'is' against the charge that any such derivations must be unsound or ethically trivial, and I show that the conclusion of each of my derivations meets the two criteria – determinacy and categoricalness – that Alan Gewirth (1979) imposes on the conclusion of any 'morally significant' derivation.

My claim, that it's possible, by means of logic alone, to derive moral conclusions from non-moral premises, may seem to fly in the face of the logical facts. After all, this volume features no fewer than three distinct formal proofs that it's impossible to derive authentically moral conclusions from non-moral premises. Pigden has proved that if 'ought' isn't treated as a logical symbol, it can't appear non-vacuously in the conclusion of a valid inference unless it appears in the premises, and the same goes for the rest of the distinctively moral vocabulary. Schurz has shown that even if we *do* treat 'ought' as a logical operator, if an 'ought' appears in the conclusion of a valid inference but not in the premises, then the predicates within the scope of the 'ought' will be *ought-irrelevant*: replaceable, without prejudice to the validity of the

inference, by any other predicates of the same grammatical type – including the opposite predicates. Hence the 'ought' in an Is/Ought inference can have no practical bite. Restall and Russell have demonstrated that in a language without the necessity operator (a significant proviso) no satisfiable set of descriptive statements entails a normative statement, where a statement is descriptive if it's preserved under normative transformations. (See Pigden 1989 and Essays 5.2 and 6.2, Schurz 1997a and Essay 6.1, and Restall and Russell, Essay 7.1.)

In other words, what all three proofs establish is that although you can derive conclusions containing moral vocabulary from premises that don't, the conclusions in question won't be substantively or authentically moral (although each proof involves a different conception of what it takes to *be* authentically or substantively moral). I accept these results, but they don't prevent me from deriving a substantively moral conclusion from premises that *aren't* substantively moral even though they *do* contain moral vocabulary. Thus my derivation of a moral conclusion from *substantively* non-moral premises is consistent with their proofs that you *can't* derive substantively moral conclusions from *formally* non-moral premises. What my derivation does tend to do, however, is deprive these proofs of philosophical significance.

What about Hume? The interpretation of Hume is of course controversial. But so far as I can see, my thesis isn't actually inconsistent with what he seems to be saying in the famous Is/Ought passage – although, again, it may lessen the passage's philosophical significance. Hume asks rhetorically for an explanation of 'what seems altogether inconceivable, how this new relation [an 'ought' conclusion] can be a deduction from others, which are entirely different from it.' Thus he seems to be implying that you can't get a conclusion in which 'ought' appears from premises in which it doesn't. Now in my derivations moral terms do appear in the premises, although not, as I'll argue, in a way that makes them moral statements. Thus on the one hand we don't get a 'new relation or affirmation' in the conclusion, since no new vocabulary is introduced. But on the other hand we do get a conclusion that is a 'deduction from [affirmations] which are entirely different from it', since the conclusion is moral while the premises aren't.

Whatever Hume may have thought, his latter-day disciples and I agree that a proposition can contain moral vocabulary without being substantively moral, a principle that's fairly obvious anyway. Consider the statement 'Some people believe that suicide is immoral'; plainly its *containing* the term 'immoral' doesn't make it a moral statement. Another, more contentious, example is premise (2) of the first derivation I give below, which (so I maintain) contains the

term 'morally wrong' while remaining substantively non-moral. Since non-moral premises can contain moral vocabulary, the formal projects of my opponents don't settle the question of the autonomy of ethics.

3. Two Presuppositions

Because I've promised to offer non-moral-to-moral derivations that are both classically *valid* and also potentially *sound*, I should address, if briefly, two metaethical positions that threaten to rule out such things from the start. First, there's the familiar view that moral utterances such as 'Torture is wrong' never express moral propositions, and so never have classical truth-values, because there *are no* moral propositions for such utterances (or any utterances) to express; call this view 'non-cognitivism'. According to non-cognitivism, no derivation of the substantively moral from premises of any sort will be classically valid, since the derivation's conclusion won't express a proposition or otherwise have a truth-value. My attempt to provide a classically valid derivation of a moral conclusion, therefore, will simply presuppose the falsity of non-cognitivism, since any attempt here to *establish* the falsity of non-cognitivism would lead us too far afield.

Second, there's the familiar view allowing that there may *exist* moral propositions but insisting that all such propositions – or at least all non-negative, atomic moral propositions – are false; call this view moral 'nihilism'. According to nihilism, no derivation of the substantively moral from premises of any sort will be sound, even if it manages to be classically valid, since the derivation's conclusion will fail to be true. (This isn't quite right, as we'll see, but it will do for now.) My attempt to provide a sound derivation of a moral conclusion, therefore, will simply presuppose the falsity of nihilism, since any attempt here to establish the falsity of nihilism would likewise be unmanageably distracting.

I recognize that both presuppositions are contentious, but I hasten to emphasize that they beg no questions in my dispute with the autonomist. The most interesting arguments for the autonomy of ethics – the arguments that deserve all the attention they have received – don't conclude that sound derivations of the substantively moral from the substantively non-moral are impossible on the grounds that substantively moral judgments are never true, either because they're neither true nor false or because they're all false. On the contrary, interesting arguments for autonomy conclude that sound derivations of the moral from the non-moral are impossible

even if moral judgments are truth-valued and even if some of those judgments are true. It's those arguments that I'm opposing here.

There's a final point to note. While, of course, no nihilist will grant that my derivations are *sound*, if either of those derivations even *validly* infers the substantively moral from the substantively non-moral, then this result will nevertheless possess considerable metaethical importance. For one thing, if there are indeed valid Is/Ought inferences, then you can't defend ethical non-naturalism on the grounds that no such inferences exist. Furthermore, you won't be able to argue *from* No-Ought-From-Is *to* non-cognitivism without begging the question by assuming non-cognitivism in order to block the Is/Ought inferences. Thus one of the chief lines of argument in favor of non-cognitivism, which starts with Hume and ends with something like emotivism, will be deprived of its dialectical power.

4. A Non-Analytic Derivation

The first derivation of the moral from the non-moral I'll call the 'Non-Analytic Derivation' since, unlike the 'Analytic Derivation' I offer in Section 7, it contains no premise that purports to be an analytic truth. The Non-Analytic version does, however, presuppose the following factual background. In 1983, Charles Rothenberg was engaged in a child-custody dispute with his ex-wife; in an attempt, as he later explained, to 'get back at' her, Rothenberg kidnapped their six-year-old son, David, and as the boy lay sleeping doused him with kerosene and set him on fire. David survived despite third-degree burns covering ninety percent of his body and remains horribly disfigured.

4.1. The Derivation

With that background in mind, consider this disjunctive syllogism:

(1) At least one (non-negative, atomic) moral proposition is true.
(2) No (non-negative, atomic) moral proposition is true, or Rothenberg's setting his son on fire was morally wrong.
Therefore:
(3) Rothenberg's setting his son on fire was morally wrong.

From (3), I take it we can deduce an appropriate 'ought'-conclusion, for instance (3*) 'Rothenberg morally ought not to have set his son on fire'. Thus, anyone tempted to object that nothing counts as a genuine Is/Ought derivation unless its conclusion contains the word 'ought' can substitute (3*) for (3) in the above derivation, modifying (2) accordingly.

4.2. Validity

Non-cognitivists to the side, most people will concede the deductive validity of the derivation: it's impossible for the premises to be true unless the conclusion is true. The derivation relies on nothing more than disjunctive syllogism, and while relevant logicians contest the validity of that rule of inference, the interest of *autonomism* as a distinct position depends on holding that valid or sound Is/Ought derivations are impossible even if we regard the classical rules of inference as valid; the autonomist, as such, doesn't blame disjunctive syllogism or anything else distinctive of classical logic for the impossibility of sound Is/Ought derivations. The derivation is also valid in a nontrivial way, assuming of course that not every application of disjunctive syllogism is trivially valid. In particular, I don't maintain that the conclusion is necessarily true and so automatically entailed by any premises at all. Nor do I maintain that either premise is impossible or that they're incompatible; on the contrary, as indicate below, I think both are true. In any event, I don't hold that (1) and (2) trivially entail (3) in virtue of entailing any conclusion at all.[2]

4.3. Soundness

The derivation also stands an excellent chance of being sound. Premise (1) must, of course, be true if any derivation has a hope of establishing the truth of a moral conclusion. To object to the derivation on the grounds that (1) is *not* true, then, is to level a rather uninteresting criticism of the derivation of 'ought' from 'is', a criticism to the effect that 'Your derivation is sound only if nihilism is false'. Of course; no one ever thought otherwise.

The other premise, (2), is highly plausible. For consider what it would take to deny it: you would have to hold that there are (non-negative, atomic) propositions that are true of the actual world, but somehow those truths don't include a proposition attributing moral wrongness to Rothenberg's setting his son on fire. I don't insist that there are *no* possible worlds of that description, only that our world isn't (likely to be) one of them. I don't maintain that (2) is necessarily, let alone analytically, true, only that (2) is highly plausible.

Nihilists (or 'error theorists', as they're also known) automatically accept (2), because they accept its first disjunct and also, presumably, the rule of disjunction-introduction. Non-nihilists of every stripe will be highly inclined to accept (2) as well. Consequentialists will find (2) plausible because they think that Rothenberg's behavior produced a net balance of unfavorable over favorable consequences, to say the least. Deontologists will find (2) plausible because they think that Rothenberg's action violated his duty not to mistreat his son or involved treating his son merely as means and not as an end in himself. But you needn't accept either nihilism or non-nihilism in order to find (2) highly plausible: you could be agnostic as between those positions while still acknowledging that 'If nihilism is true, then obviously (2) is true, and if nihilism is false, then (2) is very probably true; at worst, then, (2) is very probably true'. If, however, you don't find (2), as it stands, sufficiently plausible, substitute for its second disjunct whichever contingent proposition tops your list of the most obvious candidates for moral truth.

Premises (1) and (2) are surely not incompatible, and each of them seems plausible in its own right. Indeed, (1) and (2) are logical subcontraries, compatible propositions not both of which can be false. The less plausible you consider one of those two premises, the more plausible you must consider the other; you can't consider them both implausible.[3] Granted, again, nihilists consider their conjunction false because they consider (1) false; but, as I said before, nihilists regard the autonomy debate as uninteresting insofar as they insist that the nonexistence of sound Is/Ought derivations is already established by the truth of nihilism.

4.4. Nihilism and the Nature of Moral Propositions

Before classifying premise (1) some clarification is in order. Nihilism is often identified as the claim that all moral propositions are false, not the claim that all *non-negative*, *atomic* moral propositions are false. Thus the negation of nihilism ought to be the thesis that some moral propositions are true, not the thesis that some *non-negative*, *atomic* moral propositions are true. However, it's essential to my overall argument that premise (1) not be necessarily true and that its negation not be necessarily false. However, there's an argument that if nihilism is construed as the thesis that all moral judgments are false, it's incoherent and therefore necessarily false, which means that its negation is necessarily true. I don't regard the argument as sound, but I won't undertake to rebut it here (for criticism of the argument, see Maitzen 2006 and 2008). The argument goes as follows: If Φ is a moral proposition, then its negation ~Φ is a moral

proposition. But if Φ is false, its negation ~Φ is true. Hence not all moral propositions are false. The argument remains a worry even if we drop the assumption that the negation of a moral proposition is *always* a moral proposition, so long as we allow that in *some* cases the negation of a moral proposition is a moral proposition. For on that assumption, in those cases where the negation of a moral proposition Φ is itself a moral proposition, Φ can't be false unless its negation is true, which entails that not all moral propositions are false. What to do?

My own inclination is to stick with the traditional formulation of nihilism – that all moral propositions are false – but to fend off incoherence by adopting the following *attribution condition*:

(AC)   Any proposition Φ is a moral proposition only if Φ attributes a specific moral property to at least one thing.

Given this condition, the negation of a moral proposition will *never* be a moral proposition, and the threat of incoherence (for nihilism) and necessary truth (for its negation) evaporates. But despite my firm commitment to AC I won't be relying on it in this paper for two reasons: first, it remains controversial; second, if I adopt what some might regard as an unduly restrictive view of what counts as a moral proposition, I risk inviting the criticism that I've made things 'too easy' for myself. Let's take these points in order.

Many would object to AC on the grounds that propositions that *don't* attribute moral properties to specific things can sometimes be moral. Conditionals (including quantified conditionals with moral consequents) provide a case in point. Thus Peter Vranas (7.2) contends that 'Every citizen ought to vote' expresses a moral proposition and should be construed as a quantified conditional of the form $(x)(Cx \rightarrow O(Vx))$. I agree that 'Every citizen ought to vote' *is* a moral proposition but would argue that it's been misconstrued. It's better read as a proposition attributing the property of moral *obligatoriness* to a complex action-type, namely the action-type of voting-by-citizens (or, what comes to the same thing, as claiming that citizens and voting are two relata in the obligation relation). In that case, of course, 'Every citizen ought to vote' satisfies the attribution condition for being a moral proposition, since, in my view, action-types are things every bit as much as action-tokens are. But my opponents would reply that I'm defending an implausibly restrictive view of what counts as a moral proposition and in the course

of that defense relying on the existence of action-*types*, entities that some of them may regard with suspicion. I'd rather not raise their ontological hackles if I can avoid it. This brings me to my second reason for swallowing my convictions and shelving AC.

If a proposition is a moral proposition only if it attributes a specific moral property to at least one thing, then it follows automatically that (2) is non-moral, for the simple reason that no disjunction with distinct disjuncts qualifies as a moral proposition even if all of its disjuncts are moral. This result exposes me to the charge that I've defined my way to victory, thereby depriving my conclusions of philosophical interest: 'If by "moral proposition" you mean (in part) a proposition that attributes a specific moral property to at least one thing, then it's hardly surprising that you can conjure up inferences from the "non-moral" to the moral. Such inferences are a dime a dozen, but the reason they're so easy to produce is that you can stuff your premises with all sorts of moral vocabulary on the grounds that the propositions involved aren't really moral. Since you can help yourself to conditionals with moral consequents and disjunctions with moral disjuncts, it's child's play to derive moral conclusions from what *you* would describe as non-moral premises, to conclusions that *we* would regard as moral. But these aren't counterexamples to what *we* mean by No-Ought-From-Is. Your "Is/Ought inferences" are just Ought/Ought inferences misdescribed.' My opponents will be quick to protest that, given AC, the following qualifies as an Is/Ought or moral-to-non-moral inference:

(1b)   Brutus and Cassius's slaying of Caesar was not right.

(2*)   Either Brutus and Cassius' slaying of Caesar was right or Brutus and Cassius's slaying of Caesar was wrong.
Therefore:

(3**)   Brutus and Cassius's slaying of Caesar was wrong.

Although I believe that this derivation does (despite their protests) count as an Is/Ought inference, I recognize that not everyone will accept a criterion of moral content from which this follows. So although I will be arguing *to* something like AC, in this paper I won't be arguing *from* it.

But this brings us back to our original problem. If nihilism is the thesis that all moral claims are false, then according to some people it risks being incoherent. For if the negation of a

moral proposition is even *sometimes* a moral proposition then it can't be that all moral propositions are false. But it's crucial to my argument that something like nihilism is conceivable. In that case, nihilism must be re-described.

My concession, following a suggestion of Charles Pigden's, is to restrict the nihilistic thesis to those propositions that meet AC. Thus nihilism (for the purposes of this paper) isn't the thesis that *all* moral propositions are false – it's the thesis that all propositions attributing *specific* moral properties (or relations) to *at least one specific thing* are false. For if, as the nihilists suppose, there are no moral properties or relations for anything to possess, then all such claims will be false. Pigden has essentially the same idea when he characterizes nihilism as the thesis that all *non-negative, atomic* moral propositions are false (Pigden 2007), and it's his terminology that I've agreed to follow in this paper. Now the negation of a proposition satisfying AC doesn't *itself* satisfy AC, nor is the negation of a non-negative, atomic moral proposition itself a non-negative, atomic moral proposition. So nihilism, thus characterized, is a coherent (if false) thesis, which is what I want it to be.

4.5. The Classification of Premise (1)

Let's return to the Non-Analytic Derivation. Even supposing the derivation is sound, does it manage to derive the moral from the non-moral? Given that the conclusion, (3), looks to be a moral proposition if anything is, what about the premises? There are at least two good reasons to classify premise (1) as non-moral (for additional reasons, see Maitzen 2006, pp. 455–456, 461–463). First, it's important to note that propositions, such as (1), merely asserting the existence of truths in a given domain, even though they're *about* the domain, don't automatically belong *to* the domain. Countless examples illustrate this point. The assertion that some mathematical propositions are true, while true, isn't itself a mathematical proposition; the assertion that some universal generalizations are true, while true, isn't itself a universal generalization (although it's equivalent to the negation of one); the assertion that the domain of propositions I've never entertained contains some truths, while true, isn't a proposition I've never entertained. And so on. Furthermore, consider what it takes for us to credit someone with *knowledge* in a domain. If we ask someone to demonstrate biological knowledge, we might be impressed by the reply 'Beetles belong to the order Coleoptera' but wouldn't be impressed by the reply 'At least one biological proposition is true'. By the same token, no one shows us his or her *moral* knowledge –

knowledge in the domain of morality – with the reply 'At least one (non-negative, atomic) moral proposition is true'.

Second, the Liar-related semantic paradoxes may show that, often on pain of contradiction, statements about a domain of discourse shouldn't be counted in the domain itself. One common approach to such paradoxes notes that statements such as

> (L)   This statement is false.

generate trouble only if we allow them to refer to themselves – either directly, as in L, or indirectly, as in this pair:

> (M)   The following statement is true.
> (N)   The previous statement is false.

Suppose, however, that we block such self-reference, because we forbid a statement from belonging to the 'level', or even the 'language', of the statements to which it can refer: L can't refer to itself; M and N must belong to different levels, and higher-level statements can refer to lower-level statements but not conversely. Then we avoid the trouble these statements would otherwise cause. Some regard this solution to the semantic paradoxes as drastic and *ad hoc*, but it does suggest that logical trouble may arise if we locate non-nihilism about a domain inside the domain itself. These considerations ground at least a presumption in favor of classifying premise (1) as non-moral.

4.6. Classification of Premise (2)
Naturally, whether premise (2),

> No (non-negative, atomic) moral proposition is true, or Rothenberg's setting his
> son on fire was morally wrong,

counts as substantively moral depends on the criteria for being substantively moral. Not surprisingly, (2) tends to come out as substantively moral according to the criteria suggested –

although not explicitly stated – by the autonomists Pigden and Schurz. Of course, both would admit that a premise can contain moral vocabulary without being substantively moral. For Pigden, an expression has inference-relative vacuity in the *conclusion* of a valid inference if it can be uniformly replaced *salva validitate,* that is, without prejudice to the validity of the resulting inference. A conclusion isn't *substantively* moral if the moral vocabulary it contains is vacuous in this sense. Does it follow that a premise isn't substantively moral if the moral vocabulary it contains suffers from inference-relative vacuity? No. It's easy enough to re-jig the concept of inference-relative vacuity so that it applies to premises. An expression has inference-relative vacuity in the premises of a valid inference if it can be uniformly replaced *salva validitate*. However, a set of premises can still be substantively moral even if the moral expressions involved are vacuous in this sense. Consider, for example, Prior's first inference:

(1p)   Tea-drinking is common in England.

(2p)   Therefore either tea-drinking is common in England or all New Zealanders ought to be shot.

Here, of course, the premise is clearly non-moral since it contains no moral vocabulary whatsoever, and the conclusion isn't substantively moral either, since the 'ought' it contains suffers from inference-relative vacuity. But suppose we modify the inference thus:

(1p*)   Tea-drinking is common in England and bear-baiting is wrong.

(2p)   Therefore either tea-drinking is common in England or all New Zealanders ought to be shot.

The second conjunct in (1p*) plays no part in the derivation and could be replaced (along with everything in it) without prejudice to the validity of the inference. Nevertheless (1p*) is a substantively moral claim despite the inference-relative vacuity of its only moral constituent, since it comes out true only if bear-baiting is in fact wrong. For Pigden, therefore, a moral expression can suffer from inference-relative vacuity and still supply moral content to the premises in which it appears. If, on the other hand, the moral vocabulary is non-vacuous – and especially if it plays a part in an inference to a substantively moral conclusion – then, he would

say, this strongly suggests that the premise in which it appears is itself substantively moral. But containing non-vacuous moral vocabulary isn't, for Pigden, an infallible symptom of moral content. For example, he believes that if a moral expression occurs within the scope of an intentional operator – if somebody *believes, supposes* or *thinks* that something is good or bad, right or wrong – then the expression won't confer moral content on the premise in which it appears, even if it *does* play a non-vacuous part in an inference to a moral conclusion. Consider, for example, the following inference, M:

(1m) Chairman Mao believes that the People's Liberation Army ought to replace the peasants' doors.

(2m) The People's Liberation Army ought to do whatever Chairman Mao believes that the People's Liberation Army ought to do.

(3m) Therefore, the People's Liberation Army ought to replace the peasants' doors.

Inference M is arguably valid and one in which the 'ought' in (1m) appears non-vacuously, playing a part in an inference to a substantively moral conclusion, yet (1m) itself isn't substantively moral.

Nevertheless, Pigden would contend that if a moral expression occurs non-vacuously in the premises of a valid inference to a moral conclusion, and if it *doesn't* occur in an intentional context, then, pending strong arguments to the contrary, the premise is probably moral. Now, the expression 'wrong' in (2) is non-vacuous. Suppose we replaced the 'wrong' in (2) with 'right', yielding

(2**) No moral proposition is true, or Rothenberg's setting his son on fire was morally right.

Then of course we couldn't infer (3), that Rothenberg's setting his son on fire was morally wrong, from (1) and (2**). Hence, pending strong arguments to the contrary, Pigden would regard (2) as substantively moral.

Schurz would agree and for similar reasons. For Schurz, the conclusion of a valid inference won't be substantively moral if the moral expressions involved are ought-irrelevant with respect to the premises, that is, if the predicates occurring within their scopes can be uniformly replaced by any other predicates of the appropriate grammatical type without prejudice to the validity of the resulting inference. (Thus, for Schurz as for Pigden, substantive moral status is, at least in some cases, an inference-relative matter.) The idea of ought-irrelevance can be extended, without undue strain, to the premises of an inference. A moral expression occurs with ought-irrelevance in the premises of a valid inference if the predicates within its scope can be uniformly replaced by any other expressions of the right grammatical kind without prejudice to the validity of the resulting inference. Would Schurz say that a premise isn't substantively moral if the moral expressions involved are ought-irrelevant? I think not. Consider again, the inference adapted from Prior:

(1p*)   Tea-drinking is common in England and bear-baiting is wrong.

(2p)   Therefore either tea-drinking is common in England or all New Zealanders ought to be shot.

'Bear-baiting', which occurs within the scope of 'wrong', can be replaced with any other expression of the appropriate grammatical type – 'cock-fighting', 'being kind to animals' – without prejudice to the validity of the inference. Nevertheless, (1p*) is clearly substantively moral: first, because not all such substitutions can be made without prejudice to the *truth* of (1p*) and, second, because all such substitutions would clearly betoken a substantial change in its moral content. (It's one thing to say that *tea-drinking is common in England and bear-baiting is wrong* and quite another to say that *tea-drinking is common in England and cock-fighting is wrong*, even if both are true.) Moreover, Schurz, like Pigden, would agree that generally speaking a moral expression won't taint a premise with moral content if it occurs within the scope of an intentional operator, even if the moral expression in question *isn't* ought-irrelevant. (Thus premise (1m) of inference M is non-moral, even though we couldn't substitute 'lay down their weapons' for 'replace the peasants' doors' within the scope of the 'ought' without prejudice both to the validity of the inference and the truth of the premise.) But although the moral expressions contained in a premise can be ought-irrelevant without depriving that premise of

moral content, I think Schurz would say that if a moral expression *isn't* ought-irrelevant and *doesn't* occur within the scope of an intentional operator, then, pending powerful arguments to the contrary, the premise in which it appears is probably moral. The 'wrong' in premise (2) isn't ought-irrelevant. If we replaced (2) with

(2t)   No non-negative, atomic moral proposition is true, or Mother Theresa was morally wrong to help the poor of Calcutta,

then not only would the inference cease to be valid, but the modified second premise would cease to be true. Thus, for Schurz, pending powerful arguments to the contrary (2) comes out as substantively moral, since the 'wrong' in (2) doesn't occur within the scope of an intentional operator and is plainly *not* ought-irrelevant.

What about Russell and Restall? Taking R to symbolize 'Rothenberg set fire to his son', (2) is a disjunction of the form '$P \lor O{\sim}R$'. Thus at first sight premise (2) would appear to be disjunction with one descriptive disjunct and one moral or normative disjunct. What is the status of such propositions according to Restall and Russell? Well, if I understand them correctly their general view is that 'mixed' disjunctions, that is disjunctions with one descriptive and one normative disjunct, are *neither* normative (or moral) *nor* descriptive. Which means of course that (2) is non-moral. But there is a complication in the present case. Given their semantics, for Restall and Russell the first disjunct – no non-negative, atomic moral proposition is true – amounts to the claim that there are no ideal worlds, or that the set of ideal worlds is identical to the empty set. This claim is fragile under normative extensions. Suppose that the proposition 'no non-negative, atomic moral proposition is true' is satisfied by a model. Then there will be another model including the first, with ideal worlds added, so that the set of ideal worlds no longer coincides with the empty set. Since some non-negative moral propositions will be true at that world, the proposition that no non-negative, atomic moral proposition is true won't be satisfied by the extended model. Which means that the first disjunct is fragile under normative extensions, and consequently normative (or moral) on the Restall/Russell criterion. Since the second disjunct is also normative, the whole of premise (2) comes out as normative or moral according to the Restall/Russell account. Thus for Restall and Russell, as for Pigden and Schurz, premise (2) qualifies as a moral proposition, although for them it's essentially moral while for

Schurz and Pigden its semantic character is partly determined by its context. (We might say that in this context it *functions* as a moral proposition.)

Many autonomists would be inclined to regard (2) as moral for an additional reason. There's an intuition that, at least in many cases, the classification of a true proposition is determined by the fact or facts that make it true. Now, if my derivation is sound and thus both its premises are true, then the fact that makes (2) true is a moral fact, namely that *Rothenberg's setting his son on fire was morally wrong*. But if the fact that makes (2) true is a moral fact, then it seems as though (2) is a moral proposition.

I'll argue, however, that premise (2) is in fact non-moral. I'll argue, first, that (2) is possibly non-moral – non-moral in at least one possible world – and, second, that a proposition's classification as moral or non-moral is essential to it, in which case (2) is actually non-moral. The first claim is relatively uncontroversial; the second is much more contentious. Accordingly, the bulk of the next section is devoted to arguing that propositions possess their classification essentially, not accidentally.

Why do I say that in some possible circumstances (2) would be non-moral? Because (2) would still be true even if nothing were good, bad, right or wrong: (2) is consistent with nihilism. If nihilism were true, the fact that made (2) true would be the fact that no non-negative, atomic moral proposition is true (or that nothing instantiates any moral properties or relations), and that is about as non-moral as a fact can be. My opponents themselves contend that a disjunction derives its classification from the fact or facts that make it true. If nihilism were true, the truth-maker for (2) would be the metaethical or semantic, but at any rate non-moral, fact that no non-negative, atomic moral proposition is true. Hence the very argument of my opponents that would make (2) moral in the actual world, where nihilism is false, would make (2) non-moral in those possible worlds in which nihilism is true. Therefore those who contend that (2) is *actually* moral can hardly deny that it's possibly *non-moral*.

At this point, then, we have two options: a contingency thesis and an essentiality thesis. The contingency thesis asserts that a given proposition's classification as moral or non-moral can vary across possible worlds; the essentiality thesis, on the other hand, asserts that a proposition's status as moral or non-moral is essential to it, invariant across the possible worlds in which it exists. If we accept the essentiality thesis, then, having shown that (2) is non-moral in at least one

world, I have shown that (2) is non-moral in every world in which it exists and so non-moral in the actual world.

5. Taxonomic Essentialism

In its most plausible version, the contingency thesis threatens my project with a dilemma: either premise (1) of my derivation is actually false, or premise (2) of the derivation is actually moral. If the dilemma is genuine, then I can't maintain both that (1) is in fact true, as it must be if my derivation is sound, and that (2) is in fact non-moral, as it must be if my derivation lacks any moral premises. This dilemma depends on the contingency thesis – in this case, that the classification of a true disjunction depends on which of its disjuncts makes it true – and on the following line of reasoning: in any world in which nihilism is true, the truth-maker for (2) isn't a moral fact and so (2) is non-moral in such a world; in any world in which nihilism is false, the truth-maker for (2) *is* a moral fact and so (2) is moral in such a world.

To refute this argument I have to refute the contingency thesis at least as it applies to disjunctions. First, the contingency thesis seems to confuse meaning, or truth-*conditions*, with truth-*values*. What we mean is up to us, but whether what we mean is true is up to the world. But if what we mean is up to us – if it's determined by our communicative intentions, collective or otherwise – then surely the taxonomic status of what we mean is up to us too, since what *kind* of thing we mean depends on what it *is*, in particular, that we mean. But the contingency thesis makes the kind of thing we mean – not the wide content of our utterance (as semantic externalists already insist) but its very taxonomic *category* – depend on something besides our communicative intentions, namely, the way the world is. Thus, according to the contingency thesis, whether (2) is moral depends on whether nihilism is true. But if the kind of thing we mean is determined in part by the way the world is, then the kind of thing we mean depends not only on our intentions but also on extraneous and ostensibly nonsemantic facts. This result ought to bother those who defend the contingency thesis.

A second, and related, problem is that the contingency thesis makes us implausibly ignorant of the correct classification of disjunctions such as

(GR)   Goldbach's Conjecture is true, or Rothenberg's setting his son on fire was morally wrong,

since we don't, and perhaps can't, know the truth-value of one of the disjuncts. The contingency thesis, therefore, implies that we don't, and perhaps can't, know whether GR is moral. Possibly (if implausibly) only its first disjunct is true, in which case GR turns out mathematical and non-moral. Perhaps only its second disjunct is true, in which case GR turns out moral and nonmathematical. Perhaps, instead, the truth of GR is overdetermined by the truth of each disjunct; what is its classification then? It seems odd to say that we can't classify a proposition all of whose components we understand without first knowing which, if any, of those components makes it true. In general, we ought to be able to identify a proposition as moral or not without having to know its truth-value. On the basis of its implausible consequences, then, I reject the contingency thesis and conclude that the classification of disjunctions doesn't vary across worlds. Thus, premise (2) of my derivation, being possibly non-moral, is actually non-moral.

6. Moral Significance

We have, then, a valid and potentially sound derivation of a substantively moral conclusion from two premises each of which is substantively non-moral. The derivation is also, I believe, ethically nontrivial. Gewirth has complained that some well-known derivations of the moral from the non-moral, even if they are sound, are nevertheless not 'morally significant', since they lack two features that he says any morally significant derivation must have: determinacy and categoricalness. According to Gewirth, in order for a derivation to possess determinacy, 'the conclusion must have definite contents such that the opposed contents cannot be inferred from the same premise or premises' (Gewirth, 1979, p. 232). I doubt that we can validly derive a conclusion 'opposed' to (3) from the premises in my derivation. In order for the conclusion of a derivation to possess categoricalness,

> The requirements it sets forth must be normatively ineluctable or necessary, in that their obligatoriness cannot be removed or evaded by consulting variable, escapable features of the persons addressed or of their social relations, including their self-interested desires, their variable choices, opinions, or attitudes, and institutional rules whose obligatoriness may itself be doubtful or variable.

Unless, implausibly, Gewirth requires that every categorical conclusion be logically necessary, then (3) is categorical enough, I think, that it also meets the second of his two conditions for a morally significant derivation.

It's worth noting too that the conclusion is substantively moral, or nontrivial, in the senses suggested by Pigden, Schurz, Restall and Russell. The phrase 'morally wrong' in (3) doesn't suffer from inference-relative vacuity. If we were to replace it with 'morally right' producing

(3r)   Rothenberg's setting his son on fire was morally right,

we would no longer have either a valid argument or a true conclusion. Furthermore, what occurs within the scope of 'morally wrong' doesn't suffer from ought-irrelevance with respect to the derivation. If we replaced 'Rothenberg's setting his son on fire' with 'Rothenberg's drawing breath on his fifth birthday', to produce

(3b)   Rothenberg's drawing breath on his fifth birthday was morally wrong,

we would no longer have either a valid inference or a true conclusion. Hence the moral vocabulary in (3) is neither vacuous (in Pigden's sense) nor ought-irrelevant (in Schurz's). Thus (3) counts as substantively moral by the criteria they suggest. Finally the conclusion is normatively fragile in one of the senses suggested by Restall and Russell, since there are normative (if morally bizarre) extensions of the models satisfying (3) in which the set of all ideal worlds includes some worlds in which Rothenberg sets fire to his son.

7. An Analytic Derivation

For those unconvinced that the Non-Analytic Derivation succeeds in deriving the substantively moral from the substantively non-moral, I offer a variation whose premise (2a) has a good claim to be analytically true:

(1)   At least one (non-negative, atomic) moral proposition is true.

(2a)   No (non-negative, atomic) moral proposition is true, or any case of deliberately setting a normal, sleeping child on fire is morally wrong unless it produces a net balance of favorable over unfavorable consequences.

Therefore:

(2b)   Any case of deliberately setting a normal, sentient child on fire is morally wrong unless it produces a net balance of favorable over unfavorable consequences. [From (1) and (2a)]

(2c)   Rothenberg's setting his son on fire was a case of deliberately setting a normal, sentient child on fire without producing a net balance of favorable over unfavorable consequences.

Therefore:

(3)   Rothenberg's setting his son on fire was morally wrong. [From (2b) and (2c)]

This derivation validly employs disjunctive syllogism in deducing the intermediate conclusion (2b) from (1) and (2a); it validly employs universal instantiation in deducing the main conclusion (3) from (2b) and (2c). The only substantive additions are premises (2a) and (2c), which I will take up in reverse order.

In addition to being a tragically unfortunate truth, (2c) is a non-moral proposition. The facts that make (2c) true are partly psychological: Charles Rothenberg's action was indeed deliberate rather than inadvertent. They are partly physiological: David Rothenberg was indeed a normal, sentient (if temporarily asleep) child when his father set him on fire. It would, moreover, be the height of non-moral skepticism to suspend judgment about whether the unfavorable consequences of his being set on fire outweighed the favorable consequences. To be sure, the publicity of David's suffering briefly drew valuable attention to the issue of child-abuse, but just ask David whether that short-lived publicity was worth what he endured and continues to endure.

Premise (2a) has a good claim to be analytically true in virtue of its extremely weak second disjunct. The logic of the term 'unless' in (2a) is to be understood as it is in typical logic textbooks, where 'Φ unless Ψ' is treated as equivalent to '$\sim\Phi \supset \Psi$.' In other words, the clause 'unless it produces a net balance of favorable over unfavorable consequences' gives only a

necessary condition for the moral permissibility of deliberately setting a normal, sentient child on fire, a necessary condition that such an action may in fact never satisfy. Because '$\sim\Phi \supset \Psi$' is compatible with 'Φ & $\sim\Psi$', you can accept the second disjunct of (2a) – i.e., (2b) – even if you insist that it's *impossible* to produce net favorable consequences by deliberately setting a normal, sentient child on fire. Nihilists may reject the second disjunct in (2a) because they believe it has counterinstances, including Rothenberg's very conduct, which they regard as *not morally wrong* (and not morally right, either), regardless of its consequences. But, of course, nihilists already accept the first disjunct in (2a), since it's a statement of their position. Indeed, it seems the only way you could possess the relevant concepts and yet deny the second disjunct in (2a) would be for you to accept the first disjunct and insist that ethics is all humbug. In short, the relevant concepts make it analytically true that (1) implies (2b), in which case the disjunction $\sim$(1) v (2b) – i.e., premise (2a) – is likewise analytically true.

Now, an analytically true disjunction with contingent disjuncts can be known to be true without knowing which of its disjuncts makes it true (as Wittgenstein said, 'I know nothing of the weather when I know that it is either raining or not raining'), which strongly suggests that the classification of an analytic disjunction isn't determined by whichever fact constitutes its truth-maker. Furthermore, if analytic truths are true in virtue of the meanings of words or in virtue of conceptual connections, then there's a sense in which the truth-maker of an analytic disjunction isn't whichever disjunct happens to be true. Of course, I don't mean to imply that premise (2a) would be true if *per impossibile* both of its disjuncts were false. The point is rather that such an eventuality can be ruled out *a priori* on conceptual grounds. Now if (2a) is analytic – if its truth doesn't depend on whether *in fact* deliberately setting a normal, sentient child on fire is morally wrong – then we have good reason to suppose that it isn't substantively moral.

One might try to deny that (2a) is analytic on the grounds that we can conceive of a 'Karamazov world' (named for Dostoevsky's character Ivan Karamazov) in which 'Everything is permitted': a world in which there are true moral propositions to the effect that such-and-such is morally permissible because *every possible action* possesses the property of moral rightness, including deliberately setting a normal, sentient child on fire without producing a net balance of favorable over unfavorable consequences. I reply by challenging the conceivability, the conceptual coherence, of such a world. To my mind, the concept of moral permissibility can have no application in a world in which moral *impermissibility* could have no application.

Conceptually speaking, the moral properties come as a sort of package deal: where there can be no wrong, there can be no right. It's significant that Dostoevsky himself doesn't make this conceptual mistake: the atheistic world described by Ivan Karamazov is supposed to be a morally *nihilistic* world, a world in which nothing is morally wrong because nothing has any moral properties at all, since such properties allegedly depend on the will of God. In such a world, neither is any non-negative, atomic moral proposition true, and thus (2a) emerges as true in virtue of the truth of its first disjunct.

One might concede that (2a) is analytic but deny that its analyticity makes it non-moral. In particular, if premise (1) is true, then, whether (2a) is analytic or not, the fact that *actually* makes (2a) true is a moral fact, namely the fact stated by (2b). If the fact that makes (2a) true is a moral fact, then (2a) remains an authentically moral proposition. This line about the classification of true disjunctions should sound familiar: it's simply the version of the contingency thesis I rebutted in Section 5; my remarks in that section therefore apply here as well.

8. Post Prior

Decades ago, Prior famously used disjunctive syllogism as the second prong of a two-prong attack on the autonomy of ethics (Prior, 1960, Essay 1.1). Consider any moral proposition M and any non-moral proposition F whose negation, $\sim F$, is also non-moral. Prior's strategy was an argument by dilemma. Either $(F \lor M)$ is moral, or it's non-moral. If it's moral, then we have an inference from the non-moral premise F to the moral conclusion $(F \lor M)$. Suppose, on the other hand, that $(F \lor M)$ is non-moral. Then we have an inference from the non-moral premises $(F \lor M)$ and $\sim F$ to the moral conclusion M. Either way, we derive a moral conclusion from non-moral premises. How, then, have I offered anything new? I will conclude by comparing my strategy to Prior's.

Prior's opponents – including Pigden, Schurz, Restall and Russell, Karmo, and Humberstone – meet his dilemma by denying its initial premise. (See Essays 6.1, 6.2, 7.1, Karmo 1988, Humberstone 1982, and Humberstone 1996.) In their various ways they all deny that $(F \lor M)$ is *either* moral *or* non-moral. They all maintain that $(F \lor M)$ is *sometimes* moral (given certain assumptions and in certain contexts) and *sometimes* not (given other assumptions and in other contexts). In other words, they rely on some kind of contingency thesis. Thus, in the

inference from F to (F ∨ M), the conclusion (F ∨ M) isn't moral since the moral expressions involved will be ought-irrelevant or will suffer from inference-relative vacuity. But in the inference from (F ∨ M) and ~F to M, the premise (F ∨ M) *is* moral, since the moral expressions involved aren't ought-irrelevant and don't suffer from inference-relative vacuity. So although the conclusion is moral, it's derived partly from moral premises.

I agree with Prior and disagree with his opponents. Either (F ∨ M) is essentially moral, or it's essentially non-moral. Its classification doesn't depend on the truth-values of its components or its role within an inference: it's an absolute. But unlike Prior I get off the fence and grasp one of the horns of Prior's dilemma. (F ∨ M) is *not* moral since, like any contingent disjunction with moral and non-moral disjuncts, (F ∨ M) is compatible with nihilism, and any proposition compatible with nihilism is possibly (and so actually) a non-moral proposition. On the other hand, I agree with Prior's latter-day opponents and disagree with Prior. The inference from F to (F ∨ M) isn't, as he seems to allow, an inference from non-moral premises to a moral conclusion. But, then again, I agree with Prior and disagree with his opponents: the inference from ~F and (F ∨ M) to M *is* a valid inference from non-moral premises to a moral conclusion.

Whereas Prior uses disjunctive syllogism to establish a conditional conclusion – that *if* we classify a disjunction with one moral disjunct as non-moral, *then* we can add a non-moral premise to derive a moral conclusion from non-moral premises – I use it to establish a categorical conclusion: that *since* a particular disjunction with one moral disjunct is non-moral we can *actually* add one or more non-moral premises to derive a moral conclusion. Moreover, the Analytic Derivation contains an analytically true disjunctive premise that is, because of its analyticity, more clearly non-moral than any of the disjunctive premises Prior employs.

Furthermore, whereas I'm interested in the potential *soundness* of Is/Ought derivations, as are many of my opponents in the autonomy debate,[5] Prior was obviously interested only in validity. Consider the first premise in the only disjunctive syllogism that Prior in fact offers (Prior 1960, p. 202):

(3p) Either grass is blue or smoking is wrong.

(4p) Grass is not blue.

(5p) Therefore, smoking is wrong.

Why accept (3p)? There may be something to say in favor of its second disjunct, but there isn't much to say in favor of its first disjunct. Not even bluegrass is, I think, really blue – but if it is blue, we then lack a reason to accept (4p). By contrast, the disjunctive premises of my derivations, (2) and (2a) respectively, gain plausibility from the following fact: the less inclined you are to accept one of their disjuncts, the more inclined you should be to accept the other; indeed, in the case of (2a) rejecting one of the disjuncts analytically commits you to accepting the other. The same surely can't be said for Prior's (3p). To conclude: I believe I've produced a valid, potentially sound, and nontrivial derivation of a determinate and categorical moral conclusion from substantively non-moral premises. If I have, then the issue of the logical autonomy of ethics has indeed been settled, but not in the way ethical autonomists have supposed.

NOTES

1.See the papers by Salwén (5.1) and Pigden (5.2) for a fuller discussion of this issue. However, there's a crucial difference between my view and theirs. I contend that No-Ought-From-Is can't support an argument for non-cognitivism because it's false. They contend that No-Ought-From-Is can't support an argument for non-cognitivism even though it's true.

2. On the classical conception of deductive validity, according to which a derivation is valid if and only if it's impossible for its premises all to be true and its conclusion false: (1) any derivation with logically impossible or mutually inconsistent premises is classically valid, no matter what its conclusion; (2) any derivation whose conclusion is necessarily true is classically valid, no matter what its premises (and classically sound provided only that its premises are true). My derivation is classically valid, but not in either of those trivial ways.

3. Regarding (1) as (probably) false commits you to the (probable) truth of (2) (i.e., $\sim$(1) $\vee$ (3)), and regarding (2) as (probably) false commits you to the (probable) truth of ((1) & $\sim$(3)) and thus, of course, to the (probable) truth of (1). By contrast, however, regarding (1) as (probably) true doesn't commit you to the (probable) falsity of (2), and regarding (2) as (probably) true doesn't commit you to the (probable) falsity of (1)

4. I don't mean to imply that $\sim\Phi$ is a moral proposition whenever Φ is a moral proposition. Indeed, since I subscribe to AC, I believe the contrary: $\sim\Phi$ is *never* a moral proposition if Φ is a moral proposition. I claim, instead, that it's plausible to regard (1), in particular, as a moral proposition only if it's plausible to regard $\sim$(1) as a moral proposition. In my view, neither (1) nor $\sim$(1) is a moral proposition.

5. Including Karmo and Humberstone, *op. cit.* For criticisms, see Maitzen 1998, 2006, and 2008.